

UC Irvine

UC Irvine Previously Published Works

Title

Application of temporal streamflow descriptors in hydrologic model parameter estimation

Permalink

<https://escholarship.org/uc/item/2pw6d062>

Journal

Water Resources Research, 41(6)

ISSN

0043-1397

Authors

Shamir, E
Imam, B
Gupta, HV
[et al.](#)

Publication Date

2005-06-01

DOI

10.1029/2004WR003409

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Application of temporal streamflow descriptors in hydrologic model parameter estimation

Eylon Shamir,¹ Bisher Imam,² Hoshin V. Gupta, and Soroosh Sorooshian²

Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA.

Received 10 June 2004; revised 9 March 2005; accepted 1 April 2005; published 29 June 2005.

[1] This paper presents a parameter estimation approach based on hydrograph descriptors that capture dominant streamflow characteristics at three timescales (monthly, yearly, and record extent). The scheme, entitled hydrograph descriptors multitemporal sensitivity analyses (HYDMUS), yields an ensemble of model simulations generated from a reduced parameter space, based on a set of streamflow descriptors that emphasize the timescale dynamics of streamflow record. In this procedure the posterior distributions of model parameters derived at coarser timescales are used to sample model parameters for the next finer timescale. The procedure was used to estimate the parameters of the Sacramento soil moisture accounting model (SAC-SMA) for the Leaf River, Mississippi. The results indicated that in addition to a significant reduction in the range of parameter uncertainty, HYDMUS improved parameter identifiability for all 13 of the model parameters. The performance of the procedure was compared to four previous calibration studies on the same watershed. Although our application of HYDMUS did not explicitly consider the error at each simulation time step during the calibration process, the model performance was, in some important respects, found to be better than in previous deterministic studies.

Citation: Shamir, E., B. Imam, H. V. Gupta, and S. Sorooshian (2005), Application of temporal streamflow descriptors in hydrologic model parameter estimation, *Water Resour. Res.*, 41, W06021, doi:10.1029/2004WR003409.

1. Introduction

[2] In rainfall-runoff models, model calibration, which attempts to estimate model parameters that cannot be physically measured, is generally based on comparing simulated and measured hydrographs. Studies aiming to improve calibration methods have traditionally focused on (1) developing new optimization procedures to improve the efficiency of convergence to an optimal parameter set [e.g., Wang, 1991; Duan *et al.*, 1993], (2) selecting appropriate measures of fit (objective functions) to evaluate the model performance [e.g., Dawdy and O'Donnell, 1965; Diskin and Simon, 1977; Sorooshian and Dracup, 1980; Rao and Han, 1987], (3) utilizing probabilistic methods to account for parameter uncertainty [e.g., Kuczera, 1988; Beven and Binley, 1992; Uhlenbrook *et al.*, 1999; Thiemann *et al.*, 2001; Vrugt *et al.*, 2002], (4) calibrating the model to satisfy multiple objectives [e.g., Gupta *et al.*, 1998; Yapo *et al.*, 1998; Boyle *et al.*, 2000], and (5) developing stepwise procedures in which a subset of parameters satisfy different objectives in each step [e.g., Harlin, 1991; Hogue *et al.*, 2003].

[3] A common feature among the aforementioned procedures is their reliance on objective functions that measure performance with respect to a given timescale. For example, the Least Squares, a commonly used objective function, aggregates the residuals (observed - simulated), which are

computed for each simulation time step, to compute a performance measure corresponding to the entire calibration period. The use of the Least Squares as an objective function selects parameters that satisfy the temporal scale of the residuals (i.e., the simulated time step). Thus, in case that such an objective function is used as a single performance criterion, it is assumed that because the estimated parameters yield a plausible description of the physical processes at higher temporal resolution, they would also perform well in coarser timescales (e.g., month, season, and year).

[4] However, at different timescales, different processes, which are affected by local climatology and basin physiographic properties, tend to become more or less dominant [Beven, 1991]. For example, evapotranspiration affects the overall water balance and has a significant influence on the seasonal and annual water budget. On the other hand, infiltration, which is dominated by the intensity and duration of precipitation events (e.g., hours to days), plays a key role in partitioning storm water at the event scale. With respect to continuous process-based models, a parameter that is thought to control the infiltration rate would have some impact on the availability of soil water for evapotranspiration. Similarly, parameters affecting evapotranspiration will have an influence on soil moisture and therefore on the storm's antecedent conditions. While this accounts, in part, for the interactions among model parameters, it also lends support to the argument that parameters are scale-dependent, as demonstrated by Koren *et al.* [1999], Finnerty *et al.* [1997], and Obled *et al.* [1994]. Therefore it can be argued that model parameters must be viewed and assessed over a continuum of temporal scales that address the scale dependency of these parameters, while at the same time accounts for their interactions.

¹Now at Hydrologic Research Center, San Diego, California, USA.

²Now at Department of Civil and Environmental Engineering, University of California, Irvine, California, USA.

[5] The above argument becomes especially important with the increasing usage of hydrologic models for multiple objectives corresponding to several timescales. For example, the operational Sacramento soil moisture accounting model (SAC-SMA), which is used mainly for flood warning applications [Burnash *et al.*, 1973], has also been used for long-term predictions of flow associated with climate change impacts [e.g., Lettenmaier and Gan, 1990] and has been coupled with a land surface model to provide long-term monthly and annual water yield predictions [e.g., Lohmann *et al.*, 2004].

[6] In this paper, we present a probabilistic parameter estimation procedure that sequentially constrains the parameter space to yield posterior parameter distributions that are consistent with the uncertainties at different timescales. In other words, the parameters for the finer timescales are selected from a posterior parameter distribution that satisfies subjective performance standards for the coarser timescales. In this procedure, the convergence of model simulated streamflow toward observations is indicated by the ability to reproduce specific characteristics of the observed hydrograph. These characteristics, described using numeral streamflow descriptors, are derived for different timescales.

[7] The study was conducted for the humid Leaf River watershed, Mississippi, using the 13-parameter Sacramento soil moisture accounting model (SAC-SMA) [Burnash, 1995]. In section 2, we briefly describe the Leaf River data set and the SAC-SMA. Section 3 elaborates on the concept of streamflow descriptors and introduces the descriptors selected for the study. Section 4 introduces the sequential parameter estimation procedure (hydrograph descriptors multitemporal sensitivity analyses, HYDMUS), which is based on an evolving parameter uncertainty covariance that sequentially reduces the sampling space of the parameters. In section 5, we evaluate the performance of the estimated parameters independently and in comparison with past studies. Finally, discussion and summary are provided in sections 6 and 7, respectively.

2. Material and Methods

2.1. Data

[8] The Leaf River near Collins, Mississippi (144 km²), is a perennial headwater basin (mean daily flow 32.4 m³ s⁻¹), which is mostly covered by mixed conifer and broad leaf forest. Data for this watershed was acquired from the Office of Hydrologic Development of the National Weather Service. Forty consecutive years of data were available for this basin which includes 6-hour cumulative mean areal precipitation, mean daily flow, and estimated mean daily potential evapotranspiration. An 11-year period (WY 1952–1962) was selected for calibration, and a 26-year period (WY 1963–1988) was selected for independent validation. Additional 3-month period was used in the beginning of the calibration and validation periods as a spin-up data to establish reliable initial conditions of the model state variables. The selection of these calibration and validation periods is consistent with previous parameter estimation studies that were conducted in the Leaf River [Brazil, 1988; Duan *et al.*, 1993; Hogue *et al.*, 2003], as further discussed in section 5.

2.2. Sacramento Soil Moisture Accounting Model

[9] The SAC-SMA is a continuous conceptual, time invariant, spatially lumped model that predicts soil moisture content and streamflow hydrographs from rainfall and potential evaporation input. A detailed model description is given by Burnash [1995]. This model has been the operational streamflow forecasting model in several of the River Forecast Centers (RFC) of the National Weather Service (NWS) for nearly three decades and is currently in use for real-time flood forecasting at about 4000 forecast points in the United States [Ingram, 1996]. It has also received substantial attention from numerous research groups that use it to study modeling issues and the hydrologic responses of specific watersheds [Brazil, 1988; Sorooshian *et al.*, 1993; Gan and Biftu, 1996; Koren *et al.*, 2000; Carpenter *et al.*, 2001].

[10] Despite the physically based mechanistic conceptualization of the model's structure, its 16 parameters cannot be directly measured in the field. In this study, 13 parameters are estimated (Table 1) using parameter ranges that were defined by Brazil [1988], while the remaining three parameters are set to fixed values (SIDE, RSERV, and RIVA are 0.3, 0.0, and 0.1, respectively) as recommended by the NWS [1999]. The model was set to run in 6-hour time steps while the streamflow hydrograph simulations were aggregated to mean daily flows.

3. Hydrograph Descriptors

3.1. Concept

[11] A streamflow hydrograph represents the time and space aggregation of the hydrologic response of the contributing area and, as such, discloses information on the dominant processes that occur at the basin scale. In this study, we broadly define a streamflow descriptor as numeral information that is derived from the observed data and is thought to represent a specific characteristic of the time series.

[12] Hydrologists commonly attempt to enhance predictability by identifying unique basin signatures through data analyses. Examples of nonparametric streamflow descriptors are provided by Richter *et al.* [1997] and Poff *et al.* [1997], who proposed a wide range of scalar descriptors and discussed their potential role in characterizing the relationship between streamflow conditions and riparian flora and fauna. Their works demonstrate the usefulness of intuitive measures of flow, which are independent of statistical assumptions, but are capable of identifying signals representing unique aspects of the long-term behavior of the basin. Of particular interest to rainfall-runoff modeling is the work of Jothityangkoon *et al.* [2001] and Farmer *et al.* [2003], who derived "water balance signatures" by plotting streamflow records in three different temporal scales. They used the interannual variability of water yield, the mean monthly variation of runoff (regime curve), and the duration curve of daily flow to evaluate the level of model complexity that is required to reproduce these streamflow signatures.

3.2. Selection of Descriptors

[13] In this study we selected a group of streamflow descriptors that collectively capture the variability and the

Table 1. The SAC-SMA Parameters With Their Reasonable Range of Values in Addition to the Parameter Values Derived From Past Studies and the Minimum, Maximum, and Median Parameter Values Derived From 100 Simulations of the HYDMUS Parameter Estimation Scheme

Parameters	Description	Range ^a		SCE ^b	Brazil ^c	MACS ^d	RFC ^e	HYDMUS ^f		
		Minimum	Maximum					10%	90%	Median
UZWWM	upper zone tension water max storage (mm)	1	150	14.09	9.00	52.9	45	101.6	134.0	116.8
UZFWM	upper zone free water max storage (mm)	1	150	62.82	39.80	55.1	20	8.83	54.25	23.49
UZK	upper zone free water lateral depletion rate (day ⁻¹)	0.1	0.5	0.10	0.200	0.345	0.310	0.354	0.486	0.43
PCTIM	impervious fraction of the watershed	0	0.1	0.0	0.003	0.007	0.005	0.039	0.049	0.04
ADIMP	additional impervious area (mm)	0	0.4	0.36	0.250	0.108	0.05	0.29	0.38	0.34
ZPERC	max percolation rate	1	250	250	250	250	55	59.9	172.5	111.6
REXP	exponent of the percolation equation	1	5	2.46	4.27	4.44	2.5	3.66	4.89	4.44
LZTWM	lower zone tension water max storage (mm)	1	1000	238	240	179	120	255.9	375.9	318.2
LZFSM	lower zone free water supplementary max storage (mm)	1	1000	3.19	40	71.5	40	19.17	99.4	54.72
LZFPF	lower zone free water primary max storage (mm)	1	1000	99.80	120	142	100	75.86	197.8	137.4
LZSK	lower zone supplementary free water depletion rate (day ⁻¹)	0.01	0.25	0.019	0.20	0.042	0.06	0.032	0.139	0.074
LZPK	lower zone primary free water depletion rate (day ⁻¹)	0.0001	0.025	0.021	0.006	0.005	0.006	0.018	0.023	0.02
PFREE	fraction percolating from upper zone to the lower zone	0	0.6	0.021	0.024	0.196	0.30	0.035	0.168	0.095

^aParameter range [Brazil, 1988].

^bShuffled complex evolution, University of Arizona [Duan et al., 1993]. Parameters were taken from Thieman et al. [2001].

^cThree-stage interactive multilevel calibration [Brazil, 1988].

^dMultistep automatic calibration scheme [Hogue et al., 2002].

^eLower Mississippi River Forecast Center manual calibration from Hogue et al. [2003].

^fThe 10th, 90th, and 50th percentiles resulting from HYDMUS.

extremes in the streamflow record at different timescales. The selection of these descriptors is based on the logic presented bellow. We claim that a plausible model should have a level of flexibility that captures these properties. The following considerations were taken into account in the selection of hydrograph descriptors: (1) the descriptors are simple, concise, nonparametric, and direct (i.e., minimal data preprocessing) and (2) the descriptors are relatively independent from each other.

[14] To consider streamflow variability at different timescales, the mean daily discharge was aggregated into three timescales: the record extent (i.e., 11 years), annual flow, and monthly flow. For the record extent, the most important descriptor is the total flow, which is a nonparametric descriptor of a specific basin. For the annual flow timescale, the range between the extremes (wettest and driest years) was selected to capture the variability of annual runoff during the calibration period of record.

[15] The monthly timescale is considered to portray the annual variability of the basin's seasonal cycle. To describe this timescale, the selected descriptors should capture various characteristics of the seasonal cycle. The following discussion explains the rationale for our selection of the eight specific indices corresponding to the monthly timescale. Figure 1 represents some features of the seasonal and annual variability of monthly streamflow. Figure 1 (top) shows the monthly total flows for the period covering water years (WY) 1952–1962. This plot demonstrates the annual variability of seasonal patterns.

[16] The monthly autocorrelation plot, presented in the middle panel, indicates the time dependency that extends to the interannual and intra-annual timescales. The annual variability of the monthly flow could be partially represented by measuring the shape of the hydrograph through the monthly rising limb density (MRLD), which is defined as the ratio between the number of peaks (a peak is defined as a flow increase followed by a decrease) and the cumulative time of the rising limbs calculated from the monthly

hydrograph. Developed by Morin et al. [2002] and later studied as a streamflow descriptor by Shamir [2003] and Shamir et al. [2005], the rising limb density descriptor represents the shape and the level of smoothness of a given hydrograph, while being insensitive to the magnitude and timing of any single time step [Shamir et al., 2005]. The MRLD is expressed in frequency units (month⁻¹) while a smooth hydrograph (e.g., one peak with a long rising limb) and a noisy hydrograph will be expressed as small and high MRLD values, respectively. MRLD values can be computed for any given duration such as annual or duration of record. In this study, the MRLD value for the calibration period is considered as a hydrograph index that measures the model's ability to capture the variability of monthly flows during the period of calibration.

[17] In Figure 1 (bottom) the 40-year cumulative mean normalized runoff (solid line) is plotted against the 40-year cumulative mean normalized precipitation. The precipitation at the Leaf River seems to be relatively evenly spread over the year, which is indicated by the apparent constant slope of the cumulative precipitation curve. On the other hand, the runoff curve can be segmented within a year into three major periods with different accumulation rate: slow, fast, and slow accumulations. The first slow period in the beginning of the water year represents a “wetting” period in which precipitation saturates the basin and runoff production is relatively low. During the subsequent “wet” period the basin is relatively saturated and precipitation is transformed to runoff at higher but seemingly uniform rates. Finally, the last part of the water year, which represents a “drying” phase, coincides with seasonal recession of the hydrograph with higher potential evapotranspiration rates (summer), which cause faster depletion of soil moisture.

[18] When a monthly timescale is considered, the seasonal variability of flow (wetting wet and drying periods) can be effectively reproduced in a model that simulates the magnitudes and the variability of total monthly flows during

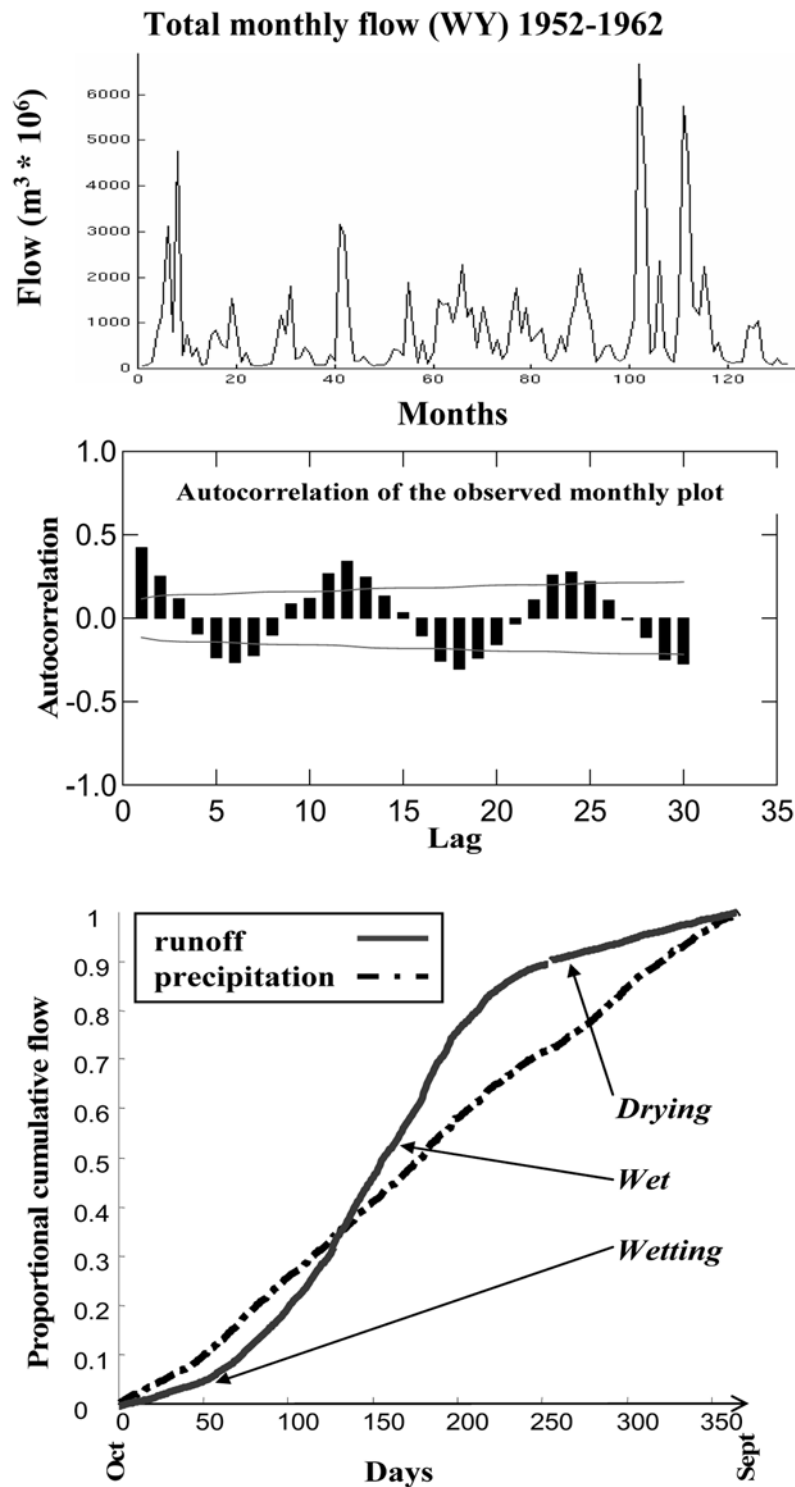


Figure 1. Characteristics of the Leaf River monthly flow: (top) Total monthly hydrograph for the WY 1952–1962, (middle) total monthly flow lag autocorrelation function, and (bottom) the 40-year mean normalized cumulative runoff and normalized cumulative precipitation.

December, April, and August. These months represent the wetting, wet and drying periods, respectively. Following evaluation of a variety of candidate descriptors, the model's ability to capture the maximum of the monthly flows during the above mentioned months was selected as indicator of its

ability to capture the annual variability of the ratio between the rainfall and runoff.

[19] The final four monthly descriptors were selected to describe the variability of flow magnitude and timing in the monthly time series. More specifically, we selected four

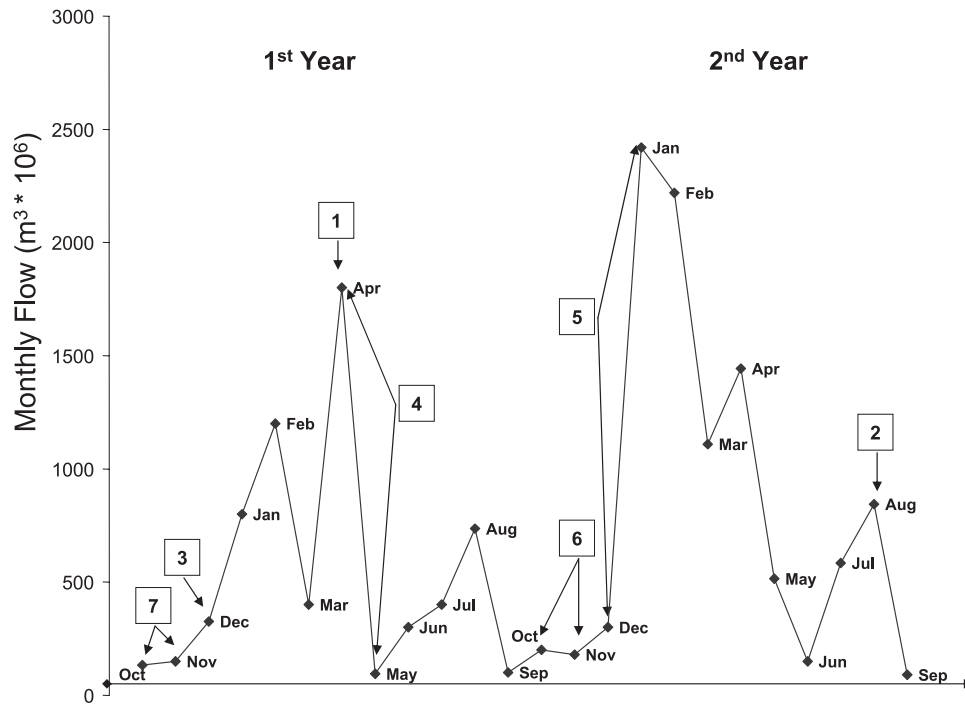


Figure 2. Hypothetical monthly flow hydrograph, which extends for 2 years and presents the eight monthly descriptors. The monthly rising limb density (MRLD) value is 0.5 month^{-1} . This value is the ratio between the number of the hydrograph peaks (Feb, Apr, Aug, Oct, Jan, Apr, and Aug, total of seven peaks) and the cumulative time during the rising limbs (Oct–Feb, Mar–Apr, May–Aug, Sep–Oct, Nov–Jan, Mar–Apr, and Jun–Aug total of 14 months). The other seven indices marked are as follows: 1–3, maximum flow occurred on April, August, and December, respectively; 4, maximum negative change; 5, maximum positive change; 6, minimum negative change; 7, minimum positive change in consecutive total monthly flow.

descriptors that indicate changes from one month to the next. Such change can be measured by magnitudes of the difference between two consecutive monthly flows. The selected descriptors are the maximum and minimum of the extreme positive and negative monthly changes (i.e., rising and recession, respectively) observed for the time series. Arguably, a model simulation that adequately captures these four extreme monthly changes should capture the monthly flow changes and variability during the observed time series. Figure 2 provides a schematic representation of the eight descriptors used in conjunction with the monthly timescale and a list of the descriptors and their associated numerical values calculated for the 11-year model calibration period (October 1952 through September 1962) is provided in Table 2.

4. Model Parameter Estimation Procedure

4.1. Procedure Development

[20] This section describes the HYDMUS scheme. The scheme uses a temporal hierarchical approach that incorporates the descriptors as model convergence measures to estimate model parameters by sequentially reducing the parameter uncertainties as the temporal resolution of the data increases. The final product of the HYDMUS scheme is the first two moments of the parameter's posterior distribution that satisfy the descriptors of the various timescales. Using a resampling procedure with the derived properties of the parameter's posterior distribution, multiple

parameter sets can be generated to provide an estimate of the effect of uncertainty due timescale dependency on the flow predictions. The rationale for the development of the HYDMUS and the procedure description is described below.

[21] Let Q_t be the observed streamflow response at a given time step t , $t = 1, \dots, n$. The hydrologic model, in our case the SAC-SMA model, can be generalized as;

$$Q_t = f(X, \theta) + \varepsilon_t \quad (1)$$

Table 2. Hydrograph Descriptors Values for the Observed Calibration Streamflow (Water Years 1952–1962)

Timescale	Descriptors ^a	Value for Water Years 1952–1962
11 years	total flow (m^3)	9.4×10^9
Yearly	annual range (m^3)	1.3×10^9
Monthly	MRLD (month^{-1})	0.58
Monthly	April maximum flow (m^3) (1)	3.6×10^8
Monthly	August maximum flow (m^3) (2)	7.4×10^7
Monthly	December maximum flow (m^3) (3)	5.0×10^8
Monthly	maximum negative change (m^3) (4)	3.9×10^8
Monthly	maximum positive change (m^3) (5)	4.8×10^8
Monthly	minimum negative change (m^3) (6)	9.6×10^5
Monthly	minimum positive change (m^3) (7)	1.6×10^5

^aThe numbers in parentheses indicate the label index of the descriptor as shown Figure 2.

where $f(\cdot)$ is the output response simulated by the model, X is the vector of input data (e.g., precipitation, potential ET) which extend from x_1, x_2, \dots, x_t , θ is the vector of model parameters to be estimated from the data; and ε_t are error terms resulting from discrepancies between the model and the real system, error in the parameters, and errors in the data.

[22] According to the paradigm of Bayesian statistical inference, the parameter vector (θ) is considered to be a random vector whose elements have a joint posterior probability density function. Initially, two sources of information about the model parameters are available: (1) the prior distribution $P(\theta)$, which is assumed by the modeler and (2) the information on model performance retrieved from comparison between model simulations (\hat{Q}) and observations $D = f(Q, \hat{Q})$. According to the Bayes formula,

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \quad (2)$$

where $P(\theta|D)$ is the posterior distribution of the parameters, $P(D|\theta)$ is the likelihood function, and $P(D)$ is a proportional constant required so that $\int P(\theta|D)d\theta = 1$ (i.e., density distribution function).

[23] The posterior distribution reflects the uncertainty in the parameters after consideration of the information (D). This is commonly done by selecting a scalar measure (objective function) y , such as the root mean square error (RMSE). In this study, the y criteria are the measures of the absolute differences between the measured and simulated 10 descriptors identified in section 3.

[24] Because of the structure complexity of hydrologic models, explicit (closed form) solution of (2) is impractical. Monte Carlo simulation methods, which are gaining popularity in hydrology [Beven and Binley, 1992; Kuczera, 1997; Thieman et al., 2001; Vrugt et al., 2003], provide a tool to deal with the inherent model complexity and the non-Gaussian multivariate distributions of the parameters [Duan et al., 1993]. Generally, in the MC simulation scheme, a prescribed number of model simulations results in an ensemble of m vectors of model predictions \hat{Q} . Each prediction (model simulation) is derived from a parameter set θ_j that is randomly drawn from the prior probability distribution $P(\theta)$. The performance of model predictions is then evaluated as to whether they capture important characteristics that are reflected in the observations. The usage of MC procedure is based on the assumption that the form of $P(\theta|D)$ can be estimated by drawing a large number of samples of $P(\theta)$, which can be used to obtain statistically significant estimates of the first two moments of the scalar (y). It is important, though, to select a sampling procedure that sufficiently samples the parameters space. In this study, model performance is measured by evaluating the ability of the model to reproduce the described streamflow descriptors with a sample size of ($m = 10,000$). This sample size was found to produce consistent results when the Monte Carlo simulations were repeated several times, indicating that the prescribed sampling size ($m = 10,000$) is adequate to represent the complex multidimensional parameter space of the SAC-SMA in the Leaf River basin.

[25] The HYDMUS procedure calls for repeated resampling (i.e., iterations) according to the number of temporal

scales to be considered. The resampling procedure enables one to collect a large sample of parameter sets and to estimate the posterior distribution associated with the different timescales. As mentioned above, in this study three temporal scales are considered (calibration period, annual, and monthly). In the first iteration, herein, the calibration period, the prior distribution can be assumed by the modeler or obtained from literature. Because no specific guidance exists regarding the prior distribution of the SAC-SMA parameters for the Leaf River Basin, we assumed a uniform distribution for the first iteration. The parameter ranges were obtained from Brazil [1988] (Table 1) and a unit covariance matrix indicating parameter independence was also assumed. The posterior distribution of the each iteration is then introduced as the prior distribution for the subsequent iteration (annual, and monthly).

[26] To derive the posterior distribution $P(\theta|D)$ of each iteration, a statistical diagnosis of the parameters' population is conducted to isolate a subset of behavioral parameters. The isolation of the behavioral parameter sets was done by selecting a subjective threshold. Following evaluations of various thresholds, we assigned thresholds that maintain good agreements between the simulated and observed descriptors and isolate ample parameter sets for meaningful parameterization of the first two moments of the posterior distributions.

[27] For the first two iterations (record extent and annual) the behavioral sets are parameter vectors that are associated with the best 10% (top 1000) of simulations as determined by the ability to capture the observed descriptor relevant to these two temporal scales (total flow and annual range, respectively). With regard to the third iteration (i.e., monthly timescale), the eight monthly descriptors were calculated from the behavioral subset of the second iteration and a 20% threshold was identified for each of the eight descriptors. The behavioral parameter sets in the third iteration are the ones that are sampled from the posterior distribution derived from the second iteration with all 8 monthly descriptors being better than the threshold defined in the second iteration. In other words, the absolute difference between the simulated and observed of each one of the 8 monthly descriptors is smaller than the absolute differences of the 20% thresholds identified in the second iteration.

[28] Once the behavioral set of parameter vectors is identified, the posterior distribution at the end of each iteration is determined by first classifying the parameter vector into sensitive and insensitive parameters. The sensitive parameters were assumed as those whose posterior distribution is significantly different from random uniform distribution. Although various tests of parameter sensitivity to the descriptors can be devised, any such test requires assignment of a subjective threshold to partition the parameters into sensitive and insensitive populations. We identified sensitive parameters as those for which the standard deviation of the normalized probability density is smaller than the threshold value of 0.12. This threshold is different from $12^{-0.5}$ which is the standard deviation of uniform distribution with range from 0 to 1, and is thought to indicate distribution that is different from uniform. Categorization to sensitive and insensitive parameters is an important step that avoids the enforcement of central tendency by

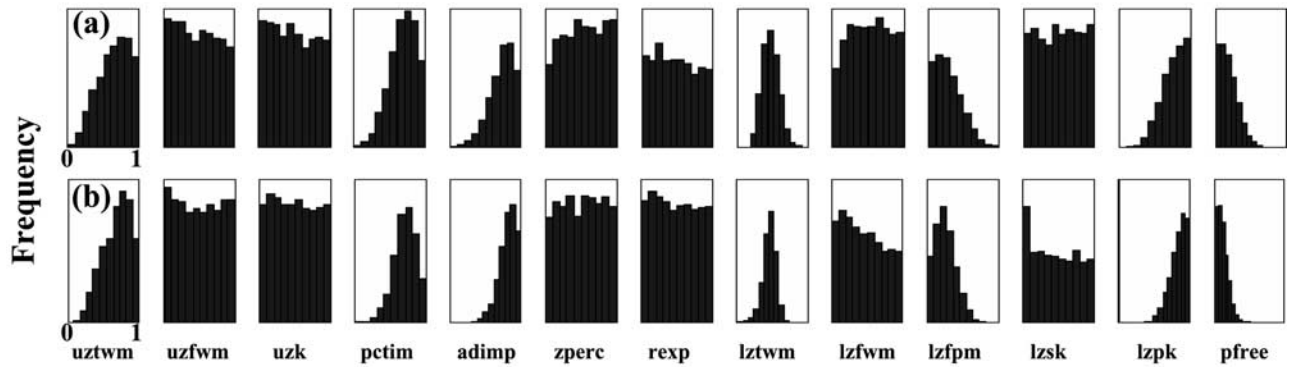


Figure 3. (a) histogram of the normalized (0–1) 1000 behavioral parameters derived from the annual selection (second iteration) and (b) resample of normalized 1000 parameters from the joint multivariate normal random sampling of the sensitive parameters and uniformly and independent of the insensitive parameters.

the resampling procedure on the posterior distribution of the insensitive parameters.

[29] Parameter resampling at subsequent timescales is accomplished in two steps: The insensitive parameters were assigned independent uniform distributions; and the sensitive parameters were assigned a multivariate normal distribution $N(\mu, \Sigma)$ with the mean vector μ and the covariance Σ estimated from the behavioral set. Such a procedure implies that the interdependency structure of the parameters could be found only among the sensitive parameters while, the insensitive parameters are independent. We assume that for a given parameter, a posterior distribution that can be approximated by the uniform distribution is an indication of insensitivity and probably also indicates independence from the other parameters.

[30] In this study, we used the mode vector as an estimator of the first moment (μ), which is considered appropriate for skewed distributions. Such commonly arises when parameters have a priori assigned parameter ranges, and the parameter distributions converge toward the extremes of their ranges. This procedure artificially fits a normal distribution to the mode using a covariance matrix that was calculated using the distributions average.

[31] To improve the resampling of sensitive parameters with a skewed distribution, parameters that are sampled outside of the a priori assigned parameter range were discarded. This practice artificially creates a skewed distribution toward the center for parameters having a mode toward the extreme ends of the range, and thereby providing a better description of the actual resulting distributions.

[32] The functioning of the resampling procedure is demonstrated in Figure 3. The behavioral parameters histograms, which are conditioned on the second iteration (i.e., annual descriptor) are compared to histograms derived from the resampling procedure (Figures 3a and 3b, respectively). It can be seen that the resampling procedure maintains key properties of the parameters distributions and produce comparable histograms. The relative similarities between the parameters distributions of the behavioral and resampled parameters were also verified using the nonparametric Kolmogorov-Smirnov test (results not shown here).

4.2. Summary of the HYDMUS Procedure

[33] A summary of the sequential steps of the HYDMUS procedure is as follows. (1) Conduct MC simulations (10,000) using uniform independent distributions of model parameters as the prior distribution for the first (coarsest) timescale. (2) Identify behavioral parameter sets by selecting parameter vectors that correspond with the best 10% of simulations as measured by the relevant descriptor. (3) Identify parameters from the behavioral parameter sets that are sensitive to the selected descriptor as those whose posterior distribution is significantly different from the uniform distribution. (4) Calculate the mode vector and covariance matrix of the identified sensitive parameters. (5) Resample 10,000 parameter sets from a joint multivariate normal distribution for the sensitive parameters and randomly and independent for the insensitive parameters. (6) Repeat steps 2–4 for the annual timescale. (7) Calculate the monthly descriptors and identify the 20% threshold for each of the eight descriptors from the behavioral parameter sets. (8) Resample from the posterior parameter distribution of the annual timescale and accept 1000 parameter sets for which all the monthly descriptors are below the 20% thresholds that were identified in step 7. (9) Derive the first two moments (mean and covariance) of the sensitive parameters from the behavioral parameter sets to represent the posterior parameter distribution.

5. Results

5.1. Parameter Uncertainty

[34] The parameters derived from the HYDMUS scheme are evaluated in this section. Figure 4 shows the 10th and 90th percentile ranges of the normalized 13 SAC-SMA parameter space, which are plotted for the three resampling iterations used in the HYDMUS. It is clear that the parameter space narrowed significantly as a result of the third HYDMUS iteration. It also can be seen that, for some parameters, the third iteration has extended the range beyond the range obtained from the previous sampling steps (UZFW, UZK, PCTIM, LZFW, and LZFP). Although the modifications of parameter ranges for these parameters are small, this suggests that different parameter

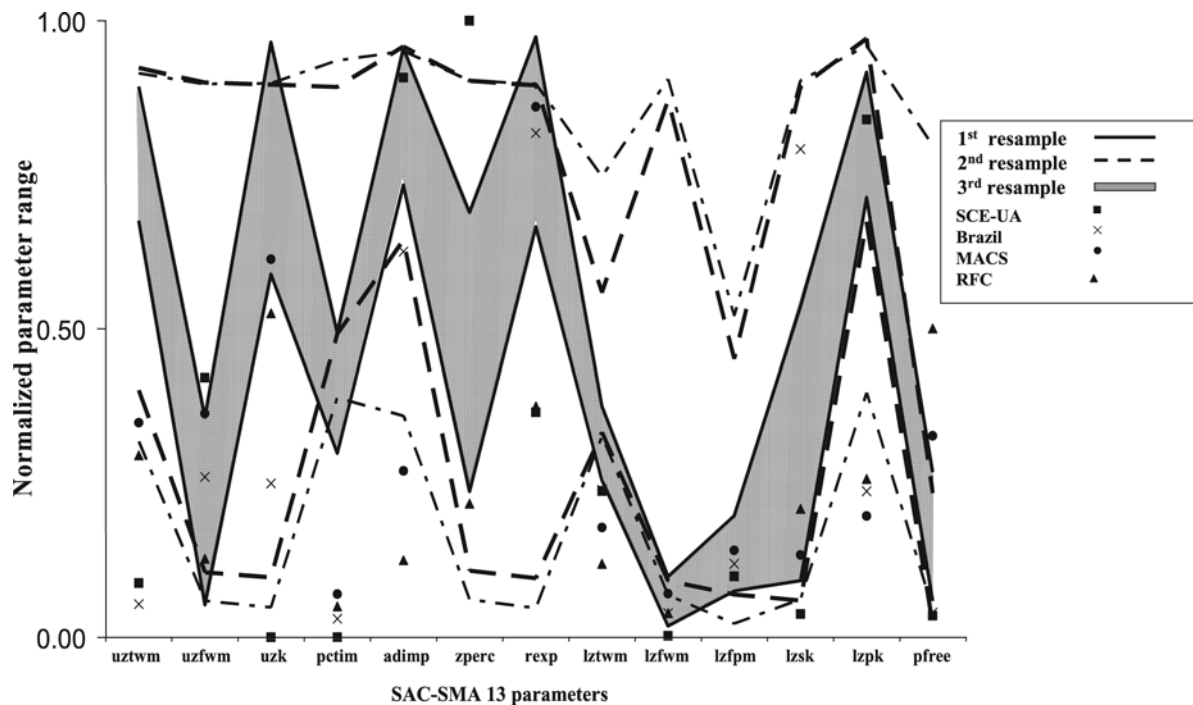


Figure 4. The 10th and 90th percentiles of the normalized SAC-SMA parameters resulting from the HYDMUS for the three resampling iterations. In addition, the normalized parameter values of the comparable past studies are shown.

values may be required to satisfy descriptors at different timescales. Also, it is an indication that the HYDMUS scheme is flexible enough to allow for shifts in the parameter space to occur.

[35] A sensitivity analysis of the parameters to the descriptors using a Monte Carlo simulation from a prior uniform distribution is shown in Figure 5. The sensitivity of a given parameter to the descriptors is identified when the behavioral cumulative distribution deviates from the cumulative uniform distribution. The parameters that show differences between the cumulative distributions are assumed to be sensitive to the descriptors. It is shown that UZFWM, UZK, ZPERC, and REXP parameters are insensitive to any of the selected descriptors. The first two parameters describe properties of the model's upper compartment, while the last two control the percolation rate. It is apparent that these four parameters are dominant in the fast hydrologic response [Bae and Georgakakos, 1994] and therefore are expected to be insensitive to coarser temporal-scale descriptors.

[36] The cumulative distributions of the resampled parameters from the three iterations are shown in Figure 6. A significant result of this study is that all of the SAC-SMA parameters were found to be identifiable in a certain range at the end of the third iteration. Note that even the parameters that are insensitive to the descriptors when selected from a uniform distribution are found to be sensitive in the third iteration. Of course, this identifiability result may be unique to the selected descriptors, and a different suite of descriptors might yield different posterior distributions.

[37] The parameter sets resulting from the HYDMUS scheme were compared to four deterministic parameter sets derived using different methods: (1) the shuffled complex evolution University of Arizona (SCE-UA) [Duan et al.,

1993], which is a global search algorithm (parameter values taken from Thiemann et al. [2001]), (2) Brazil's three-stage interactive multilevel calibration procedure [Brazil, 1988], (3) a Multistep automatic calibration scheme (MACS), which is a sequential calibration that applies different objective functions to groups of parameters sequentially [Hogue et al., 2003], and (4) a parameter set provided by the Lower Mississippi River Forecast Center personnel (hereinafter RFC), who calibrated the model using the NWS manual calibration technique [NWS, 1999] (parameter values were taken from Hogue et al. [2003]).

[38] These past studies represent techniques with various levels of automation and expert evaluation ranging from a completely automatic calibration procedure (SCE-UA) to manual calibration (RFC). The calibration in these studies was done using the simulation of the mean daily flow to identify a single optimal parameter vector. The first three studies (SCE-UA, Brazil, and MACS) used the same calibration data set as used in this study (WY 1952–1962). We recognize that since the past studies were developed to accommodate different calibration objectives and used various measure of fits criteria, a quantitative comparison among the studies is questionable. However, a comparison among the simulations of these studies that evaluate the model's ability to capture a variety of hydrologic response phenomena provides a better perspective on the HYDMUS performance.

[39] The SAC-SMA parameter values for the different studies are summarized in Table 1. In addition, the normalized parameter values from the previous studies were compared to the HYDMUS 10th to 90th percentiles parameter ranges obtained after the third iteration (see Figure 4). Clearly, the majority of the parameters derived from the HYDMUS are different from those estimated by the previ-

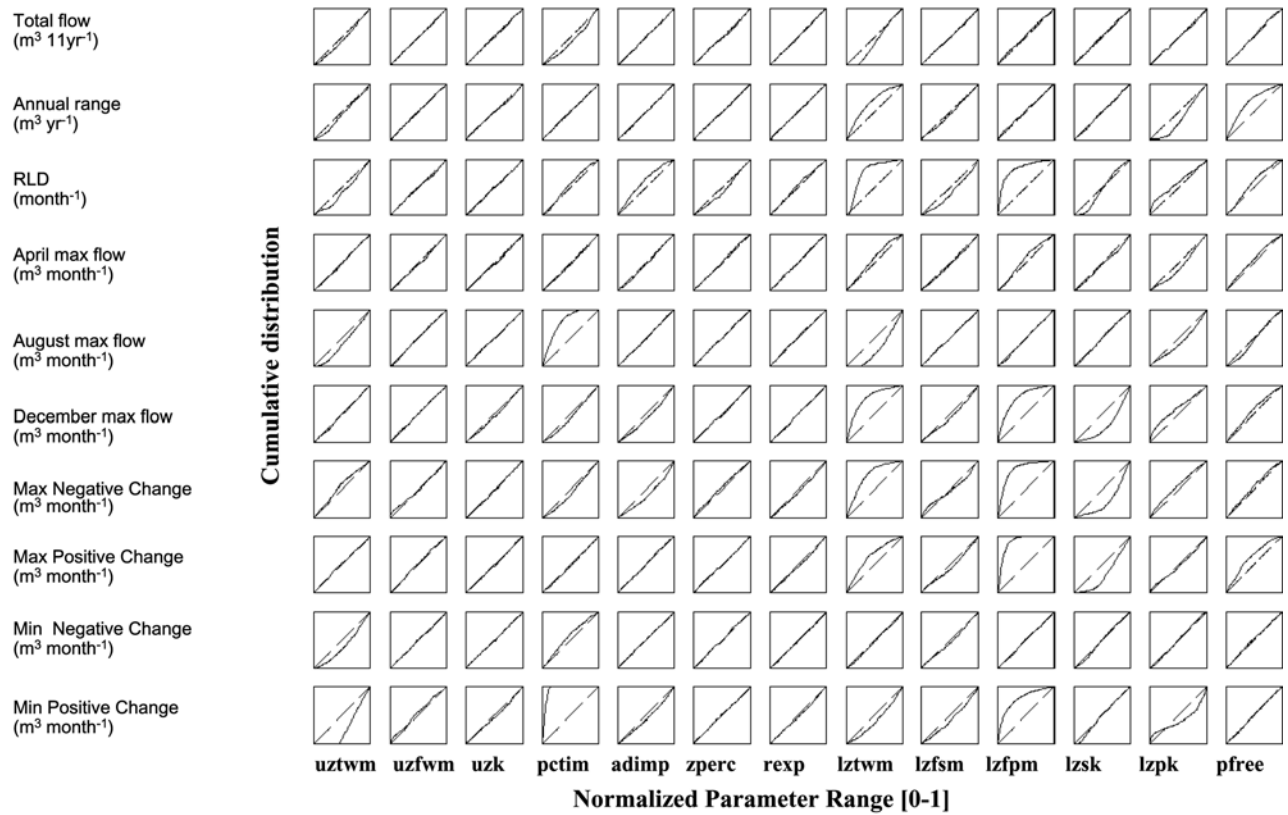


Figure 5. Sensitivity analysis of the SAC-SMA parameters to the 10 selected hydrograph descriptors. The solid line represents the cumulative distribution of the 1% behavioral parameter derived from a priori uniform parameter distribution, and the dashed line represents the cumulative distribution of the uniform distribution.

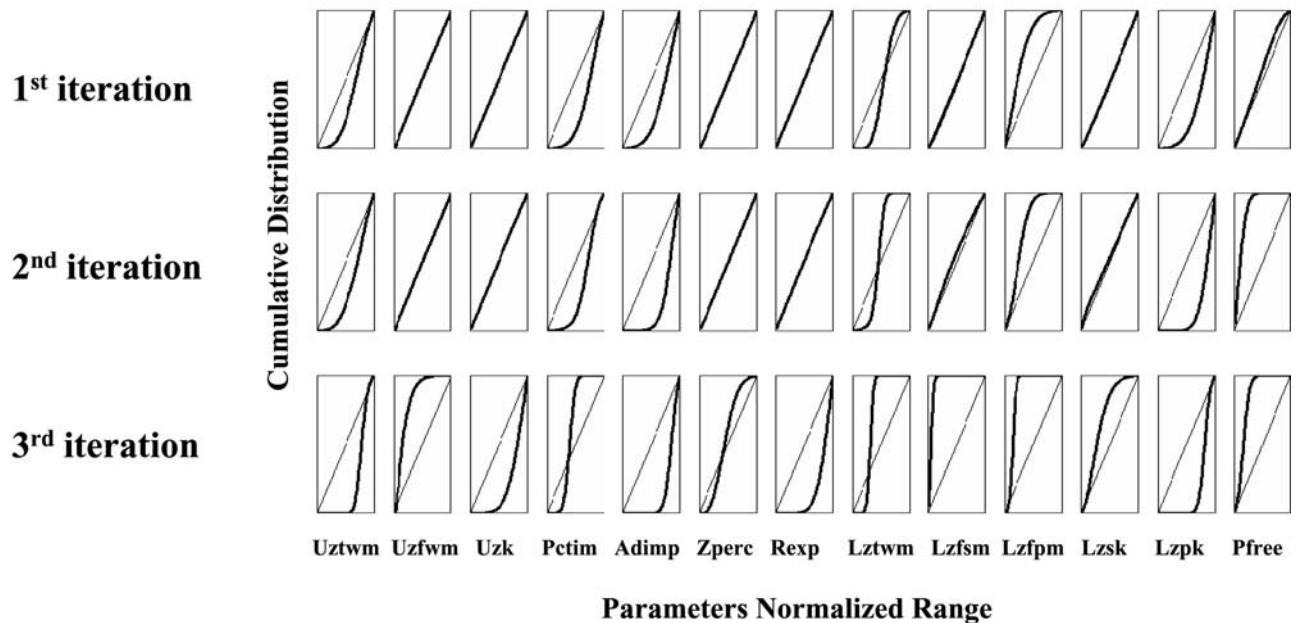


Figure 6. The cumulative distribution of the best 1% Monte Carlo behavioral simulations (thick line) and the cumulative distribution of a uniform random sample for the three resampled iterations.

Table 3. Scalar Evaluation Criteria for the Leaf River (Water Years 1963–1988) From HYDMUS Parameters and Past Four Deterministic Studies

Evaluation Criteria	Equations	SCE	Brazil	MACS	RFC	HYDMUS		
						Minimum	Maximum	Median
Total bias	$\sum_{t=1}^{n=9498} Q_{sim,t} - \sum_{t=1}^{n=9498} Q_{obs,t}, t = days,$	4576	5395	4449	6753	1596	3550	2298
Mean absolute bias (annual)	$\sum_{t=1}^{n=26} Q_{sim,t} - Q_{obs,t} , t = years$	1885	2151	1804	2595	1255	1592	1353
Mean absolute bias (monthly)	$\sum_{t=1}^{n=312} Q_{sim,t} - Q_{obs,t} , t = months$	236	246	230	286	212	326	248
Mean absolute bias (daily)	$\sum_{t=1}^{n=9498} Q_{sim,t} - Q_{obs,t} , t = days$	11	11	11	12	11	20	13
DRMS	$\sqrt{\frac{1}{n} \sum_{t=1}^{n=9498} (Q_{sim,t} - Q_{obs,t})^2}, t = days$	20	22	23	23	20	35	28
Rcoeff	$\frac{\sigma_{Q_{obs,t}Q_{sim,t}}}{\sigma_{Q_{obs,t}}\sigma_{Q_{sim,t}}}, t = days$	0.95	0.95	0.94	0.94	0.92	0.95	0.94

ous studies. For example, the SCE-UA has only three parameters that are within the range of the HYDMUS (ADIMP, LZFP, and LZPK). On the other hand, all of the previous studies were out of the range derived from the HYDMUS for the UZTWM, PCTIM, ZPERC, LZTWM, and PFREE parameters. The HYDMUS procedure convergence to a new region in the parameter space demonstrates the dependence of the selected parameters on the parameter search methods and the objective functions used in the calibration.

[40] In summary, the parameter uncertainty derived from HYDMUS provides a relatively narrow parameter range, which does not include the parameters from the previous studies. In the next section, the performance of the model using HYDMUS-derived parameters is compared with streamflow observations and the simulations from the previous studies.

5.2. Evaluation of Flow Performance

[41] A performance evaluation of the simulated flow was conducted at the three timescales considered in the application of HYDMUS (i.e., record extent, annual, monthly), in addition to evaluation of the daily timescale. The evaluation was conducted on an independent 26-year period (WY 1963–1988). This evaluation of the parameters on a longer historical time frame enables testing of the overall model performance and detection of model divergence (i.e., good performance during calibration but poor performance during the evaluation period).

[42] A suite of statistical evaluation criteria was calculated to quantify model performance (Table 3). The first four measures in Table 3 are used to evaluate the water balance at different timescales, in addition to the commonly measured daily root mean square error and correlation coefficient. Values are shown for the maximum (worst), minimum (best), and median performance derived from 100 simulations randomly resampled from the HYDMUS posterior parameter distribution.

[43] A very significant improvement of the overall total bias (flow extent) and the overall annual biases was achieved using the HYDMUS procedure as compared

with past studies. This, of course, is expected as a result of considering these two time steps in the HYDMUS procedure. It can be seen that at these timescales, even the worst value (maximum value in Table 3) derived from the HYDMUS is better than all of the past four studies. In the monthly time steps, the HYDMUS minimum value is better than the other studies and, for the daily time step, the minimum value performed better than the RFC and similar to the MACS parameters. Also, the DRMS and the correlation coefficient performance are comparable in quality to the past studies, with the minimum HYDMUS DRMS being inferior only to the SCE-UA and the correlation coefficient being as good as the SCE-UA and Brazil.

[44] Visual evaluation of the simulated annual yield performance (Figure 7) reveals that in comparison with the observations, out of the 26 years, all past studies overestimated the annual flow of at least 23 years. The two years with the largest flow volume were underestimated by all of the studies except from the RFC. It can be seen from Figure 7 that in general, the resulting HYDMUS annual volumes from the 100 simulations in most of the years is closer to the 1:1 line.

[45] The monthly hydrographs of the measured and the HYDMUS simulations for WY 1963–1965 are presented in Figure 8. The monthly flow confirming good performance of the HYDMUS simulations and capturing the seasonal dynamics. In Figure 9 the monthly flow duration curve for the 26 evaluation years is shown. The flows were transformed using a Box-Cox transformation [Box and Cox, 1964] to relate the error variance at each time step to its associated flow magnitude and to provide flow values with an approximately Gaussian distribution. The Box and Cox transformation is expressed as

$$q_{transformed} = \frac{(q_t + 1)^\lambda - 1}{\lambda} \quad (3)$$

where λ is a scale parameter selected to have the value of 0.3, based on recommendations by Misirli *et al.* [2003]. The

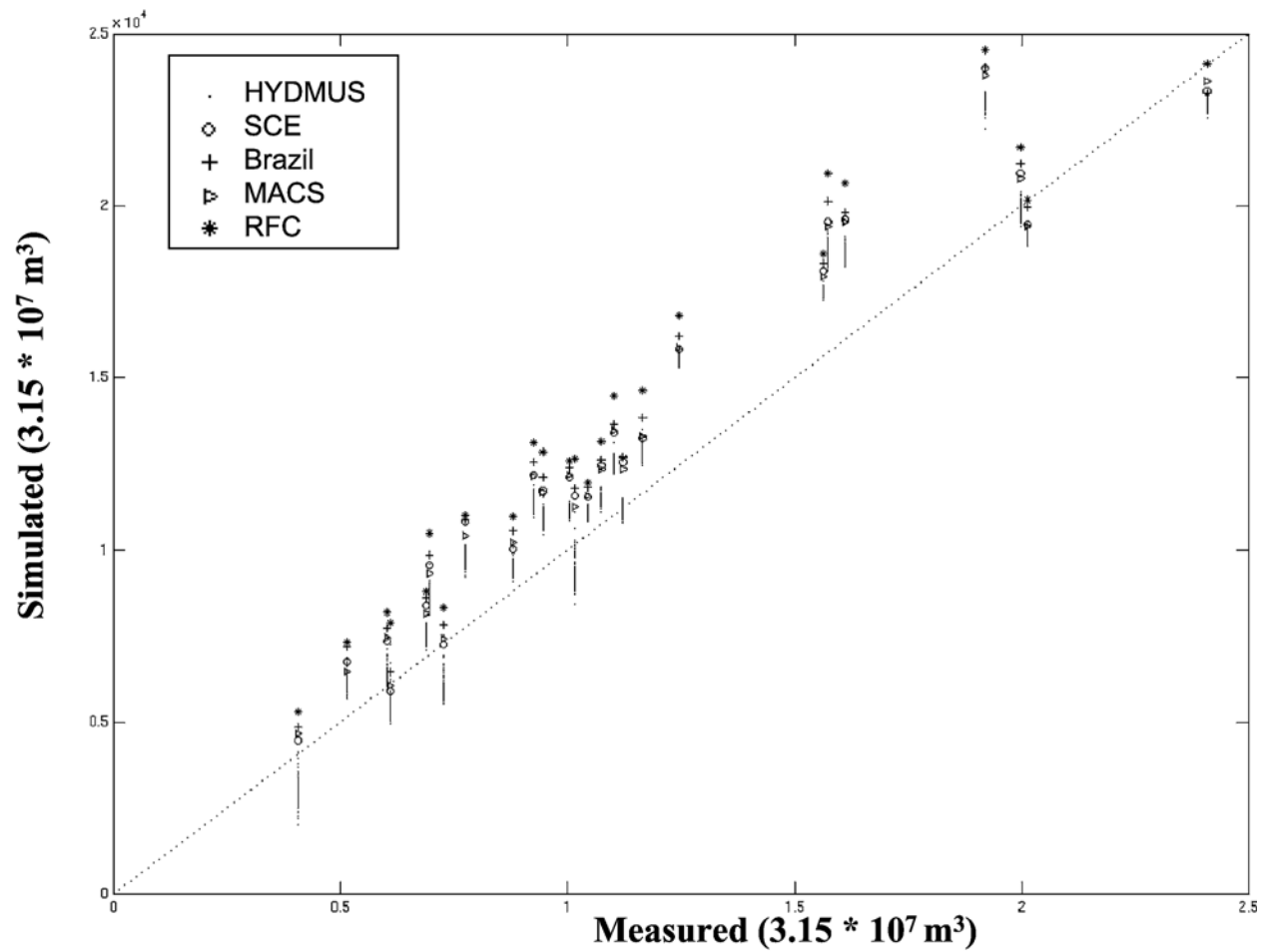


Figure 7. Simulated annual yield of 26 years (WY 1963–1988) resulting from 100 simulations of the HYDMUS posterior distribution and the four deterministic past studies compared with the observed annual flows.

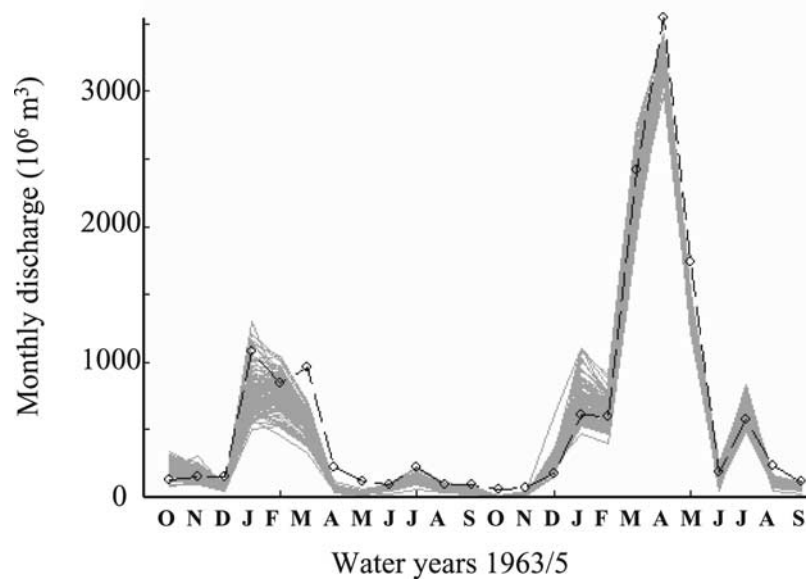


Figure 8. The monthly hydrograph (WY 1963–1965) observed flow and the prediction uncertainty derived from 100 simulations from the HYDMUS posterior distribution.

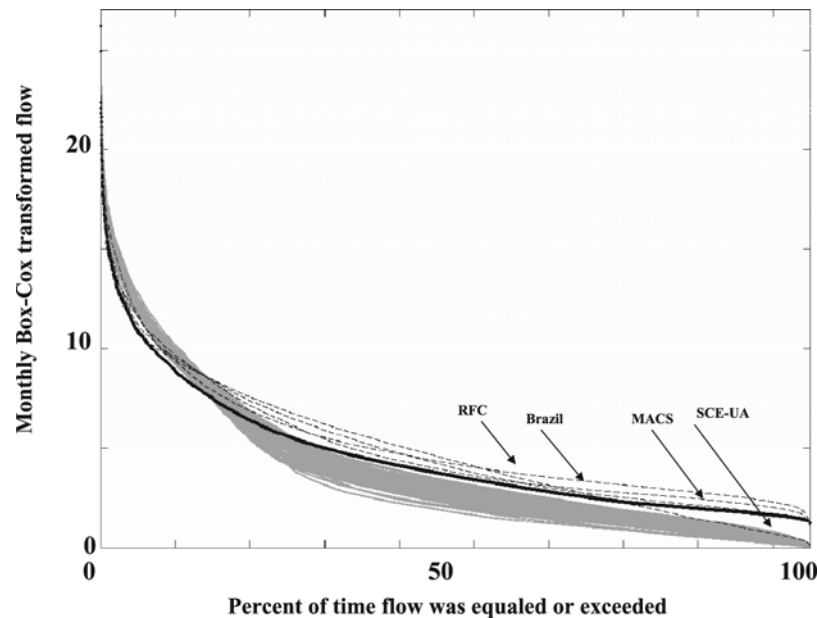


Figure 9. The monthly duration curves of the Box-Cox transformed observed (solid thick line), 100 simulations from the HYDMUS posterior distribution (shaded lines), and the past four deterministic studies (dashed lines).

monthly flow duration curve (Figure 9) indicates that the performance of the four deterministic studies is similar and, except for about 10% of the highest monthly flows, they consistently overestimate the observed monthly flow. One exception is the SCE-UA, which underestimates the lowest 80–100% exceedance of monthly flow. The ensemble of flows obtained using the HYDMUS procedure captures the highest 5% observed monthly flow events and the medium flow magnitude that is between 15–30% exceedance. Overall, it is observed that the HYDMUS overestimates and underestimates the highest and lowest monthly flows, respectively.

[46] Finally, the mean daily flow, which was not considered as a criterion in the HYDMUS calibration, was evaluated. Figures 10a and 10b present the measured and the HYDMUS simulated mean daily flow ($\text{m}^3 \text{s}^{-1}$) and Box-Cox transformed flow in the 1964 water year, respectively. From Figure 10a it can be seen that the HYDMUS simulation captures the dynamics and magnitude of the wet season flow events; however, the HYDMUS parameters are underestimating the low-flow events mainly in the drier seasons (Figure 10b).

[47] Further exploration of the daily flow is provided in Figure 11, which shows the duration curve of the daily Box-Cox transformed flow for the evaluation data set. In this curve, the measured, the HYDMUS ensemble simulations, and the aforementioned past studies are shown. Again, as also inferred from the monthly duration curve (Figure 9), the deterministic studies shown in Figure 11 are consistently overestimating the flows, except from the highest-flow events. One exception is the SCE-UA, which underestimates the low-flow events. The HYDMUS captured the high-flow events (0–5% exceedance) and the middle flow range of about 40–85% exceedance; however, it overestimated the 10–40% and underestimates the low flow (85–100%).

[48] Prediction of high-flow episodic events is clearly a performance that is required from a flood prediction model such as SAC-SMA. Although the HYDMUS procedure was not explicitly calibrated to meet this objective, the results for high-flow events can also be examined. In Figure 12, 26 observed daily flow events that exceeded $500 \text{ m}^3 \text{s}^{-1}$ are plotted with their counterpart simulations from the HYDMUS and the past studies. In Figure 12 the dots represent the median, and the error bar represents the maximum and minimum flow values resulting of 50 simulations from the posterior parameter distribution of the third iteration. It is evident that the ranges yielded from the HYDMUS contain the values resulting from the deterministic studies. In some cases, some results from the HYDMUS perform better than the deterministic study, such as in the case of the two extreme events.

[49] In summary, it appears that the HYDMUS procedure performed better than the four deterministic procedures at the yearly and record extent timescales. At the monthly timescale, using HYDMUS, results in overestimating the high flows and underestimating the low flows, whereas the previous studies consistently overestimated the monthly flow. The HYDMUS procedure results in parameters that, in general, provide less systematic (random, less biased) residuals compared with the past studies, which mostly tend to overestimate the flow. This may be because the past studies give higher priority to the high-magnitude events, which resulted in an overestimation at lower magnitude time steps. For daily timescales, the HYDMUS performance is comparable to the other studies in all of the flow regimes and, in some respects, even performs better (for instance when capturing the middle-flow magnitude).

[50] The previous four studies have an overall similar behavior, while the HYDMUS procedure produced a different set of parameters and streamflow behavior. This is clearly a result of the new descriptor-based sequential

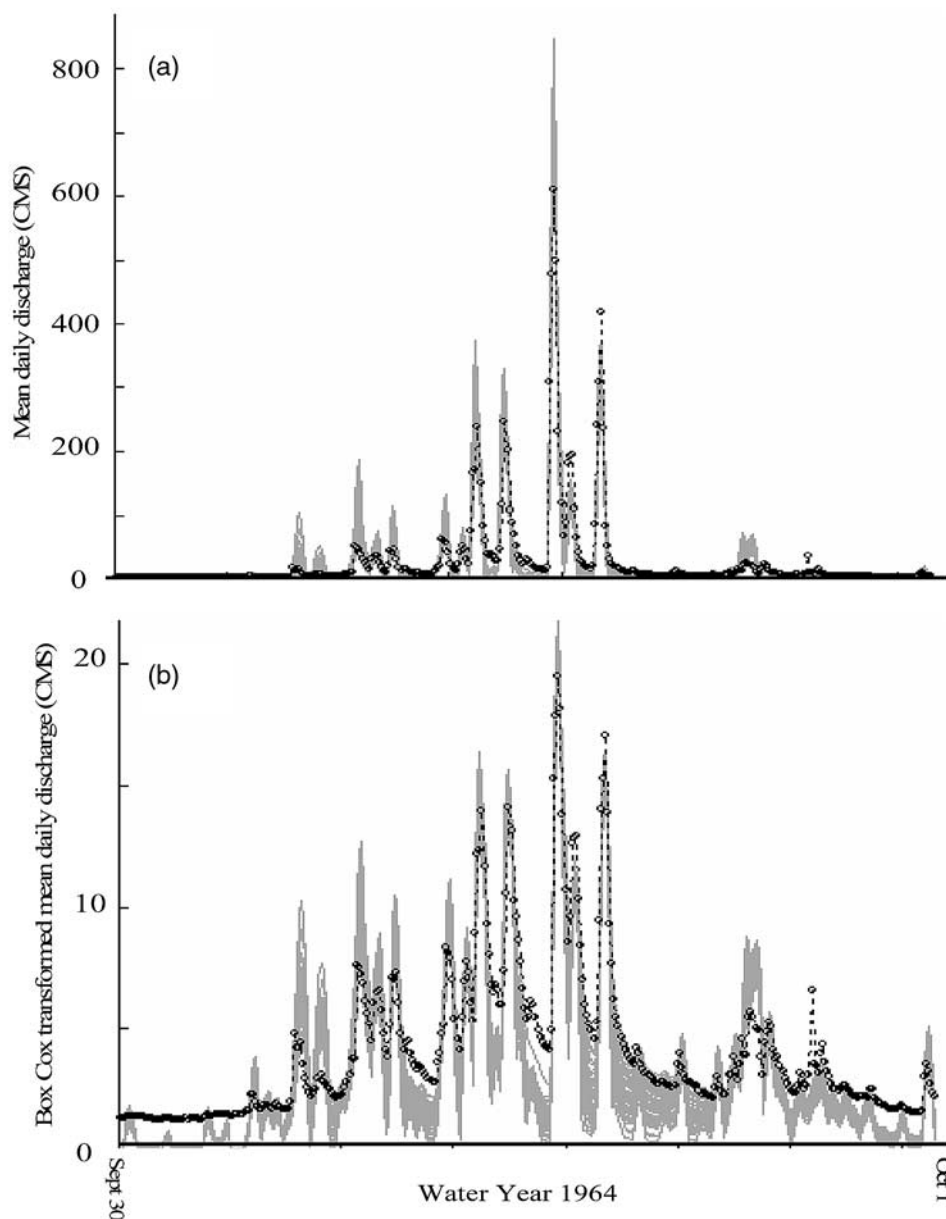


Figure 10. The daily hydrograph (WY 1964) observed flow (circles) and the prediction uncertainty derived from the HYDMUS scheme (shaded lines). (a) Flow (in $\text{m}^3 \text{s}^{-1}$) and (b) the Box-Cox transformed flow.

approach, which accentuates an exploration of new acceptable regions in the parameter space.

6. Discussion

[51] The presented HYDMUS method is a parameter estimation scheme that requires user interaction in a set of relevant decisions such as selection of the descriptors, defining the priorities of the model objectives, and selection of the threshold values in the procedure. The identification of hydrograph descriptors within the HYDMUS procedure requires analyses and subjective decisions that prioritize the timescale and the important streamflow responses. This action consequently incorpo-

rates some of the qualities of manual calibration into the automatic procedure.

[52] The derivation of descriptors from streamflow, which compared to other fluxes (e.g., precipitation, evapotranspiration) is a hydrologic flux that is relatively less prone to measurement errors and uncertainty associated with time and space interpolation, provides reliable information on the system response. Such descriptors, if consistent (i.e., repeatable with small dispersion when analyzed in various time periods) and distinguishable (i.e., unique for a given basin), can capture intrinsic physiographic characteristics of basins. In addition, descriptors that are well chosen (i.e., consistent and distinguishable) can decrease the dependency of the parameter estimation on the data sets selected for the

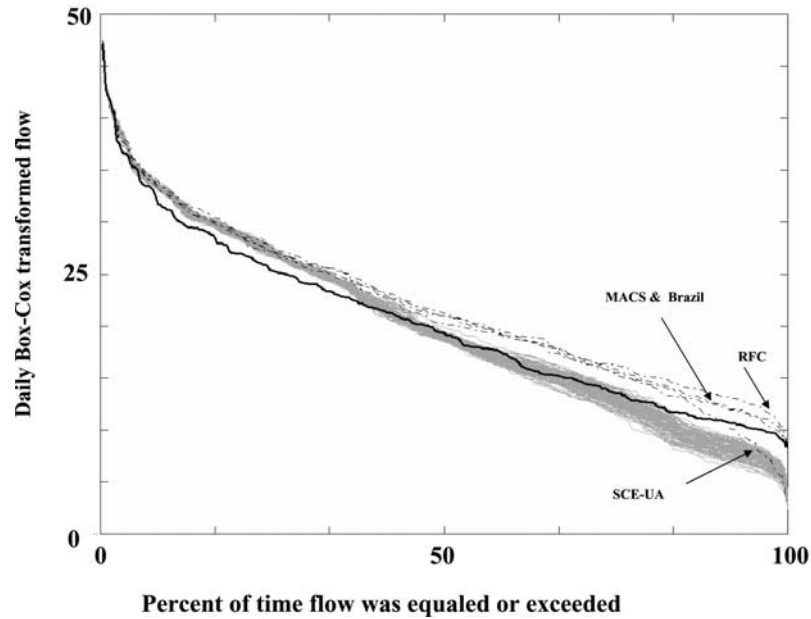


Figure 11. The daily duration curve of the Box-Cox transformed flow for 26 years (WY 1963–1988) of the observed (thick solid line), 100 simulations from the HYDMUS posterior distribution (shaded lines), and the past deterministic studies (dashed lines).

calibration. An extensive discussion and demonstration of the concept of descriptors is provided by Shamir [2003]. In this paper, we focused on defining simple descriptors that capture the dynamics and the variability of the stream that are believed to summarize important information about the dominant processes.

[53] At various timescales, driven by the climatology and basin physical properties, different dominant processes are expressed in the basin hydrological cycle. For example,

evapotranspiration is a process that affects the overall water balance and thus is dominant at seasonal and annual time-scales, while infiltration is dominant in the episodic time-scale (e.g., hourly, daily). However, hydrologic models are often constructed to describe the mechanistic response of the basin to an episodic event (e.g., a precipitation event for flood warning). Furthermore, the calibration using the traditional objective functions treats the residuals at the timescale of the simulation. Although these conceptual

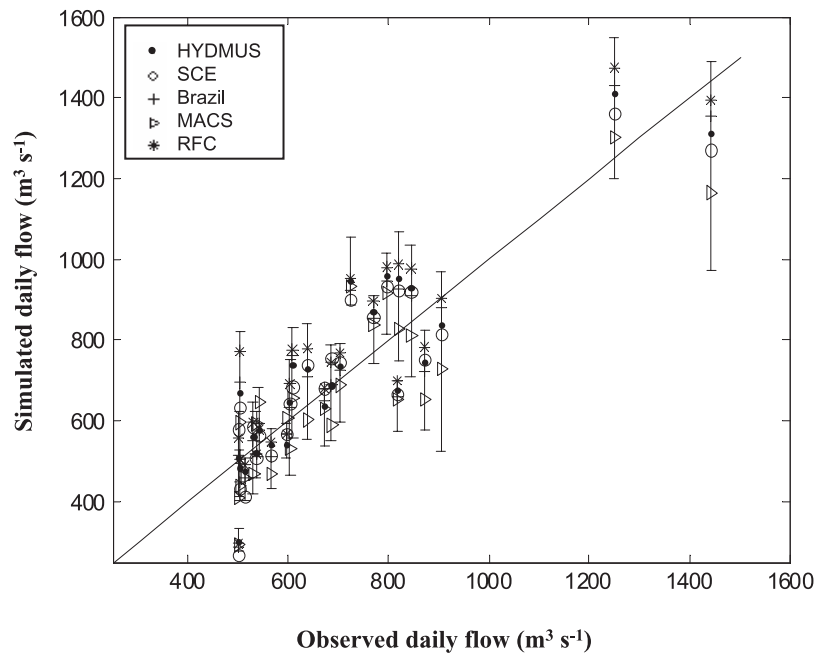


Figure 12. Daily flows that are greater than observed $500 \text{ m}^3 \text{ s}^{-1}$ (26 events in the evaluation data sets (WY 1963–1988)). The black dots represent the median, and the error bars represent the maximum and minimum flow values resulting from 50 simulations of the HYDMUS third iteration.

models spatially lump the processes to an integrated response, in the temporal sense, these models describe the episodic timescale mechanism. It is reasonable to state that a robust model, even if constructed to describe episodic events, should be required to perform well at different timescales.

[54] In this study, to assure model performance in multiple temporal scales, we selected a sequential timescale calibration method. In other words, selected model parameters at a finer timescale are estimated from the posterior distribution developed from parameters that satisfy a coarser temporal scale. It was found that, even without consideration of the daily timescale in the HYDMUS procedure, the performance at the daily time step is comparable to deterministic methods. The parameters selected from the HYDMUS procedure appear different from those provided by the other deterministic studies, indicating that the multidimensional parameter space is complex and may have different subregions of attraction that satisfy the hydrologic response in different ways. Note, however, that the parameter regions derived using HYDMUS could be different if a selection of different timescales, descriptors, and resampling sequence was made.

[55] It should be mentioned that this procedure does not contradict existing parameter estimation procedures. In fact, the practice of selecting descriptors as a measure of fit can be incorporated to complement other parameter estimation procedures, such as the generalized likelihood uncertainty estimation (GLUE) [Beven and Binley, 1992].

[56] The HYDMUS procedure also provides a test for parameter identifiability and its role in different timescales. For example, in the SAC-SMA model, the PCTIM, ADIMP, and UZTWM parameters clearly represent the response characteristics of the upper soil surface. Consequently, these parameters affect the fast response of the system to an episodic event. In the HYDMUS procedure, it was seen that the PCTIM is already sensitive at the coarser timescale (time extent), and the UZTWM and the ADIMP are sensitive at the monthly timescales (Figure 5). This sensitive behavior is explained by the level of control in the evapotranspiration processes which affect the long-term water balance. On the other hand, UZFWM and UZK, parameters that characterize the upper zone drainage rates, were found to be insensitive to the longer timescale.

7. Summary and Conclusions

[57] This paper has presented the results of a parameter estimation approach based on hydrograph analyses to identify distinct nonparametric descriptors that are characteristic of the time series. The descriptors were selected to measure important streamflow characteristics at three timescales (monthly, annual, and the record extent). A parameter estimation scheme HYDMUS (hydrograph descriptors multitemporal sensitivity analyses) was developed based on Monte Carlo sampling applied in a stepwise approach, so that parameter estimation at shorter timescales were constrained by the estimated posterior distribution of the behavioral parameters at the coarser timescales. The final product of the HYDMUS scheme is an ensemble of simulations that indicates the parametric uncertainty associated with the selection of the descriptors and the calibration timescales.

[58] The HYDMUS scheme was tested using the SAC-SMA model applied to the Leaf River, Mississippi. The procedure resulted in a relatively narrow parameter range for the 13 parameters. When evaluated on monthly, annual, and flow extent, the results were significantly better than four parameter sets derived from previous studies. At the daily time step, the model performances using parameters derived from the HYDMUS procedure were found to be comparable to the deterministic studies in the middle flow magnitudes, with a tendency to overestimate the high flows and underestimate the low flows.

[59] We believe that the hydrograph descriptor approach is a powerful way to obtain consistent parameter estimates for watershed models, and future research should attempt to explore the relationships between hydrograph descriptors and physically measurable characteristics of a watershed. Success in defining such links would potentially aid in watershed modeling for ungauged basins.

[60] **Acknowledgments.** Primary financial support for this study was provided by NASA grant NAG5-8503 (HyDIS), NASA-Raytheon grant 300623 (Synergy), SAHRA (NSF-STC for "Sustainability of semi-Arid Hydrology and Riparian Areas"), and agreement EAR-9876800. The first author is grateful for the Salt River Doctoral Fellowship in 2002–2003 and the Graduate College Fellowship for fall 2003. Thanks are extended to Konstantine Georgakakos, Efrat Morin, Thorsten Wagener, Terri Houge, Murugesu Sivapalan, Charles Kroll, and two anonymous reviewers for their valuable suggestions and critique and to Corrie Thies for proofreading the manuscript.

References

- Bae, D.-H., and K. P. Georgakakos (1994), Climatic variability of soil water in the American Midwest: part 1. Hydrologic modeling, *J. Hydrol.*, **162**, 355–377.
- Beven, K. J. (1991), Scale consideration, in *Recent Advances in Modeling of Hydrologic Systems*, edited by D. S. Bowles and P. E. O'Connell, pp. 357–371, Springer, New York.
- Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and predictive uncertainty, *Hydrol. Processes*, **6**, 279–298.
- Box, G. E. P., and O. R. Cox (1964), The analysis of transformation, *J. R. Stat. Soc., Ser. B*, **26**(2), 211–252.
- Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, **36**(12), 3663–3674.
- Brazil, L. D. (1988), Multilevel calibration strategy for complex hydrologic simulation model, Ph.D. dissertation, Dep. of Civ. Eng., Colo. State Univ., Fort Collins.
- Burnash, R. J. C. (1995), The NWS river forecast system—Catchment model, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, chap. 10, pp. 311–366, Water Resour. Publ., Highlands Ranch, Colo.
- Burnash, R. J. C., E. R. L. Ferral, and R. A. McGuire (1973), A generalized streamflow simulation system, *Rep. 220*, 204 pp., Joint Fed.-State River Forecast Cent., Sacramento, Calif.
- Carpenter, T. M., K. P. Georgakakos, and J. A. Sperflage (2001), On the parametric and NEXRAD-radar sensitivities of a distributed hydrologic model suitable for operational use, *J. Hydrol.*, **253**, 169–193.
- Dawdy, D. R., and T. O'Donnell (1965), Mathematical models of catchment behavior, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, **91**(HY4), 123–139.
- Diskin, M. H., and E. Simon (1977), A procedure for the selection of objective function for hydrologic simulation models, *J. Hydrol.*, **34**, 129–149.
- Duan, Q. Y., V. K. Gupta, and S. Sorooshian (1993), Shuffled complex evolution approach for effective and efficient global minimization, *J. Optim. Theor. Appl.*, **76**(3), 501–521.
- Farmer, D., M. Sivapalan, and C. Jothityangkoon (2003), Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis, *Water Resour. Res.*, **39**(2), 1035, doi:10.1029/2001WR000328.

- Finnerty, B. D., M. B. Smith, D.-J. Seo, V. Koren, and E. Moglen (1997), Space-time scale sensitivity of the Sacramento model to radar-gage precipitation inputs, *J. Hydrol.*, **203**, 21–38.
- Gan, T. Y., and G. F. Biftu (1996), Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure, *Water Resour. Res.*, **32**(12), 3513–3524.
- Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measure of information, *Water Resour. Res.*, **34**(4), 751–763.
- Harlin, J. (1991), Development of a process oriented calibration scheme for the HBV hydrologic model, *Nord. Hydrol.*, **22**, 15–36.
- Hogue, T. S., H. Gupta, S. Sorooshian, and C. D. Tomkins (2003), A multi-step automatic calibration scheme for watershed models, in *Calibration of Watershed Models*, *Water Sci. Appl.*, vol. 6, edited by Q. Duan et al., 346 pp., AGU, Washington, D. C.
- Ingram, J. (1996), Lessons taught by floods in the United States of America, paper presented at International Council for Science SC/IDNDR Workshop on River Flood Disasters, Int. Council. for Sci., Koblenz, Germany, 26–28 Nov.
- Jothityangkoon, C., M. Sivapalan, and D. L. Farmer (2001), Process controls of water balance variability in a large semi-arid catchment: Downward approach to hydrological model development, *J. Hydrol.*, **254**, 174–198.
- Koren, V. I., B. D. Finnerty, J. C. Schaake, M. B. Smith, D.-J. Seo, and Q.-Y. Duan (1999), Scale dependencies of hydrologic models to spatial variability of precipitation, *J. Hydrol.*, **217**, 285–302.
- Koren, V., M. Smith, D. Wang, and Z. Zhang (2000), Use of soil property in the derivation of conceptual rainfall-runoff model parameters, paper presented at the 80th Annual Meeting, Am. Meteorol. Soc., Long Beach, Calif., 9–13 Jan.
- Kuczera, G. (1988), On the validity of first-order prediction limits for conceptual hydrologic models, *J. Hydrol.*, **103**, 229–247.
- Kuczera, G. (1997), Efficient subspace probabilistic parameter optimization for catchment models, *Water Resour. Res.*, **33**(1), 177–185.
- Lettenmaier, D. P., and T. Y. Gan (1990), Hydrologic sensitivities of the Sacramento-San Joaquin River Basin, California, to global warming, *Water Resour. Res.*, **26**(1), 69–86.
- Lohmann, D., et al. (2004), Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project, *J. Geophys. Res.*, **109**, D07S91, doi:10.1029/2003JD003517.
- Misirli, F., H. V. Gupta, S. Sorooshian, and M. Thieman (2003), Bayesian recursive estimation of parameter and output uncertainty for watershed models, in *Calibration of Watershed Models*, *Water Sci. Appl.*, vol. 6, edited by Q. Duan et al., pp. 113–124, AGU, Washington, D. C.
- Morin, E., K. P. Georgakakos, U. Shamir, R. Garti, and Y. Enzel (2002), Objective, observations-based, automatic estimation of the catchment response timescale, *Water Resour. Res.*, **38**(10), 1212, doi:10.1029/2001WR000808.
- National Weather Service (NWS) (1999), National Weather Service River Forecast System (NWSRFS) user manual, report, NOAA, Silver Spring, Md.
- Obled, C., J. Wending, and K. Beven (1994), The sensitivity of hydrological models to spatial rainfall patterns: An evaluation using observed data, *J. Hydrol.*, **159**, 305–333.
- Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegard, B. D. Richter, R. E. Sparks, and J. C. Stomberg (1997), The natural flow regime, *BioScience*, **47**, 769–784.
- Rao, R., and J. Han (1987), Analysis of objective functions used in urban runoff models, *Adv. Water Resour.*, **10**, 205–211.
- Richter, B. D., J. V. Baumgartner, R. Wigington, and D. P. Braun (1997), How much water does a river need?, *Freshwater Biol.*, **37**, 231–249.
- Shamir, E. (2003), The use of streamflow indices in hydrologic modeling, Ph.D. dissertation, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tucson.
- Shamir, E., B. Imam, E. Morin, H. V. Gupta, and S. Sorooshian (2005), The role of hydrograph indices in parameter estimation of rainfall-runoff models, *Hydrol. Processes*, in press.
- Sorooshian, S., and J. A. Dracup (1980), Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour. Res.*, **16**(2), 430–442.
- Sorooshian, S., Q. Duan, and V. K. Gupta (1993), Calibration of rainfall-runoff models: Application of global optimization to the Sacramento soil moisture accounting model, *Water Resour. Res.*, **29**(4), 1185–1194.
- Thieman, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, **37**(10), 2521–2535.
- Uhlenbrook, S., J. Seibert, C. Leibundgut, and A. Rodhe (1999), Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameter and structure, *Hydrol. Sci. J.*, **44**, 779–797.
- Vrugt, J. A., W. Bouten, H. V. Gupta, and S. Sorooshian (2002), Toward improved identifiability of hydrologic model parameters: The information content of experimental data, *Water Resour. Res.*, **38**(12), 1312, doi:10.1029/2001WR001118.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, **39**(8), 1201, doi:10.1029/2002WR001642.
- Wang, Q. J. (1991), The genetic algorithm and its application to calibrating conceptual rainfall-runoff models, *Water Resour. Res.*, **27**(9), 2467–2471.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *J. Hydrol.*, **204**, 83–97.

H. V. Gupta, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA.

B. Imam and S. Sorooshian, Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA.

E. Shamir, Hydrologic Research Center, 12780 High Bluff Drive, Suite 250, San Diego, CA 92130, USA. (eshamir@hrc-lab.org)